remote sensing

MDPI

Article

Adaptive Month Matching: A Phenological Alignment Method for Transfer Learning in Cropland Segmentation

Reza Maleki [1], Falin Wu [1,*], Guoxin Qu [2], Amel Oubara [1], Loghman Fathollahi [3] and Gongliu Yang [4]

[1] SNARS Laboratory, School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China; maleki@buaa.edu.cn (R.M.); oubara@buaa.edu.cn (A.O.)
[2] Beijing System Design Institute of Electro-Mechanic Engineering, Beijing 100854, China
[3] Meteorological Department of West Azerbaijan Province, Iran Meteorological Organization (IRIMO), Orumiyeh 670056, Iran; loghman.fathollahi@student.kit.edu
[4] School of Mechanical Engineering, Zhejiang University, Hangzhou 310030, China
* Correspondence: falin.wu@buaa.edu.cn; Tel.: +86-10-82313929

Abstract: The increasing demand for food and rapid population growth have made advanced crop monitoring essential for sustainable agriculture. Deep learning models leveraging multispectral satellite imagery, like Sentinel-2, provide valuable solutions. However, transferring these models to diverse regions is challenging due to phenological differences in crop growth stages between training and target areas. This study proposes the Adaptive Month Matching (AMM) method to align the phenological stages of crops between training and target areas for enhanced transfer learning in cropland segmentation. In the AMM method, an optimal Sentinel-2 monthly time series is identified in the training area based on deep learning model performance for major crops common to both areas. A month-matching process then selects the optimal Sentinel-2 time series for the target area by aligning the phenological stages between the training and target areas. In this study, the training area covered part of the Mississippi River Delta, while the target areas included diverse regions across the US and Canada. The evaluation focused on major crops, including corn, soybeans, rice, and double-cropped winter wheat/soybeans. The trained deep learning model was transferred to the target areas, and accuracy metrics were compared across different time series chosen by various phenological alignment methods. The AMM method consistently demonstrated strong performance, particularly in transferring to rice-growing regions, achieving an overall accuracy of 98%. It often matched or exceeded other phenological matching techniques in corn segmentation, with an average overall accuracy across all target areas exceeding 79% for cropland segmentation.

Keywords: cropland segmentation; crop phenological alignment; transfer deep learning; Sentinel-2; time series

check for updates

Academic Editor: Ruyin Cao

Received: 9 November 2024
Revised: 9 January 2025
Accepted: 13 January 2025
Published: 15 January 2025

Citation: Maleki, R.; Wu, F.; Qu, G.; Oubara, A.; Fathollahi, L.; Yang, G. Adaptive Month Matching: A Phenological Alignment Method for Transfer Learning in Cropland Segmentation. *Remote Sens.* **2025**, *17*, 283. https://doi.org/10.3390/rs17020283

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The global pursuit of food security and sustainable development is becoming increasingly urgent due to growing food demands and rapid population expansion [1,2]. Advanced crop monitoring and segmentation technologies, particularly those utilizing deep learning (DL) [3–5] for analyzing multispectral satellite imagery like Sentinel-2 [2,5], have emerged as crucial tools in transforming agricultural productivity [6,7]. These technologies significantly improve land use efficiency and promote sustainable farming practices, playing a vital role in achieving the United Nations Sustainable Development Goals (SDGs) [8,9]. By ensuring the sustainability of food production systems, these advancements contribute

to the well-being of future generations [9–11]. Moreover, integrating DL with satellite imagery facilitates precise crop classification and monitoring, essential for optimizing agricultural output and resource management [12].

In regions without advanced agricultural monitoring systems like the United States' Cropland Data Layer (CDL) [13,14], managing crops and ensuring food security presents significant challenges. The availability of open-source satellite data, particularly Sentinel-2 imagery, accessible using platforms such as Google Earth Engine (GEE) [2,5,15,16], is crucial for bridging this data gap, providing detailed agricultural insights essential for sustainable farming practices. Transfer learning stands out as a powerful strategy [17], enabling the use of models trained on extensive datasets to improve agricultural monitoring in regions with limited data [4,14,18]. This approach highlights the global significance of satellite imagery in advancing agricultural productivity and sustainability. By leveraging transfer learning, we can enhance the precision of crop monitoring and classification in data-scarce areas, thereby contributing to significant improvements in global food security [17,19].

Despite advancements in DL and satellite imagery for agricultural monitoring, several challenges remain. Current methodologies often rely heavily on local cropland maps and specific regional data, limiting their applicability in diverse and data-scarce regions [20]. Additionally, the variability in crop phenology [21] across different geographical areas necessitates adaptive methods that can accurately align and interpret satellite data for effective crop classification and segmentation [12,22]. Existing approaches, such as the CDL in the United States, do not readily extend to regions lacking detailed agricultural datasets, creating a significant gap in global food security efforts [14,23]. Traditional methods may not fully leverage the temporal dynamics of crop growth, resulting in suboptimal segmentation accuracy. These issues underscore the need for advanced, adaptive methods to enhance the transferability and robustness of crop monitoring models across varying agricultural landscapes [18,24].

In agricultural monitoring using remote sensing, the debate on data utilization focuses on balancing the need for comprehensive crop cycle coverage with targeted analysis [25]. Some researchers advocate for annual datasets to capture full crop cycles, while others prefer focusing on specific growing seasons to improve image quality and reduce atmospheric interference [26]. However, in transfer learning for DL models, the alignment of training data with the target area's major crops is often overlooked. For example, Ge et al. [23] primarily focused on corn and rice, neglecting other key crops like soybeans, which are crucial for improving model accuracy. Optimizing time series based on the phenological stages of the target area's crops is essential for better segmentation performance [22,27].

Also, phenological stages are critical for crop identification and segmentation in remote sensing, as crops display distinct characteristics during specific growth periods. For instance, paddy rice is best identified during its irrigation phase, while corn and soybeans are more distinguishable at the flowering or heading stages [16]. While these stages are generally stable, the exact timing can vary due to factors such as crop variety, topography, soil, climate, and farming practices across different regions [14,27]. This variability complicates the transfer of crop models between areas with different growth patterns.

For phenological alignment, Sakamoto et al. [28] proposed the shape model fitting (SMF) method, which aligns crop phenology using a vegetation index (VI) time series shape model through linear shift and scaling operations. The SMF method employs parameters such as time shift ($tshift$), time scaling ($xscale$), and VI scaling ($yscale$) to minimize the root mean square error (RMSE) between observed and modeled VI curves. The SMF method employs a shape model, denoted as $g(\mathrm{x})$, representing a VI time series linked to ground

phenology observations, to align it with the seasonal VI time series of the same crop type, referred to as $h(\mathrm{x})$, through linear shifting and scaling operations, described as

$$\hat{g}(\mathrm{x}) = yscale \times [g(xscale \times (x + tshift)) + bias] - bias. \quad (1)$$

Here, *bias* represents a fixed value specific to each crop type, determined by the background VI value. The parameters *xscale* and *yscale* serve as scaling factors for the two dimensions (time and VI value). Using the parameters *xscale* and $tshift$, the phenological stages of $h(\mathrm{x})$ can be calculated based on their corresponding stages in $g(\mathrm{x})$:

$$p^i_{est} = \frac{1}{xscale} \times p^i_0 - tshift,\ i = 1, \ldots n \quad (2)$$

where $h(\mathrm{x})$ is represented as p^i_{est} and $g(x)$ as p^i_0 for $i = 1, \ldots n$.

While effective, SMF has inherent limitations, including a global matching strategy that applies uniform scaling across all phenological stages, leading to inaccuracies for early-stage phenology estimates. Subsequent improvements, such as the localized SMF-S [29] and the spatiotemporal SMF (STSMF) method [30], addressed some of these issues by segmenting phenological stages or incorporating spatiotemporal neighboring pixel information. However, these methods require parameter optimization, which increases computational cost and complexity, especially when applied to large-scale datasets with diverse phenological patterns. Additionally, methods like STSMF demand data on phenological stages from prior years, limiting their scalability. These challenges highlight the need for a more adaptive and efficient alignment method.

Transfer learning in crop segmentation has attempted to address these challenges through various strategies. One common approach is to enhance classifiers for better transferability, though generalizability remains a concern [31]. Another method involves using multitemporal datasets over several years to account for variability in crop growth due to environmental differences [32], though this can be resource-intensive. Some studies have integrated multiple satellite sources, such as Sentinel-1 and Sentinel-2, to improve temporal resolution and capture phenological changes [15,16]. Despite these efforts, aligning temporal characteristics between training and target areas remains a key challenge in transfer learning [23].

This study addressed the challenges in transferring advanced crop monitoring technologies to diverse and data-scarce regions by proposing an Adaptive Month Matching (AMM) method. This method optimizes time series data based on the phenological stages of major crops in the target area, such as corn, soybeans, rice, and double-cropped winter wheat/soybeans (Dbl-Crop), thereby enhancing the transferability and accuracy of DL models. The AMM method aligns Normalized Difference Vegetation Index (NDVI) profiles between training and target areas to account for phenological differences. Additionally, we systematically evaluated the performance of the AMM method by comparing it with other techniques and the best-matched time series using known cropland information, employing visual interpretation as an additional comparative method and assessing the results against other research to validate the robustness and effectiveness of the AMM method.

2. Materials and Methods

2.1. Study Area

The study area consisted of a designated training area and several target areas across diverse regions of the United States and Canada, as shown on Figure 1. The training area is located on some parts of the Mississippi River Delta in Arkansas, Tennessee, Missouri,

and small parts of Kentucky and Mississippi. This place has unique seasonal phenological phases and a diverse agricultural landscape, making it ideal for developing robust crop classification and segmentation models [4,23,33,34]. The target areas included T10SEH, T14TNK, T14TNR, T15TUG, T16SCF, T16SCJ, T16TFM, and T17TNJ. Each target area is designated by its UTM (Universal Transverse Mercator) zone and specific tiles according to the MGRS (Military Grid Reference System). For instance, T14TNK refers to a tile within UTM Zone 14, and T10SEH refers to tile SEH within UTM Zone 10. These target areas were selected to represent a variety of crop types and environmental conditions, allowing for a comprehensive evaluation of the methods across different agricultural landscapes. Additionally, T17TNJ (Canada) was chosen as a target area to evaluate the transferability of the methods beyond the USA. This area was selected for its similarity in crop types to the Mississippi River Delta (training area), including corn, soybeans, and Dbl-Crop, making it suitable for evaluating transfer learning.

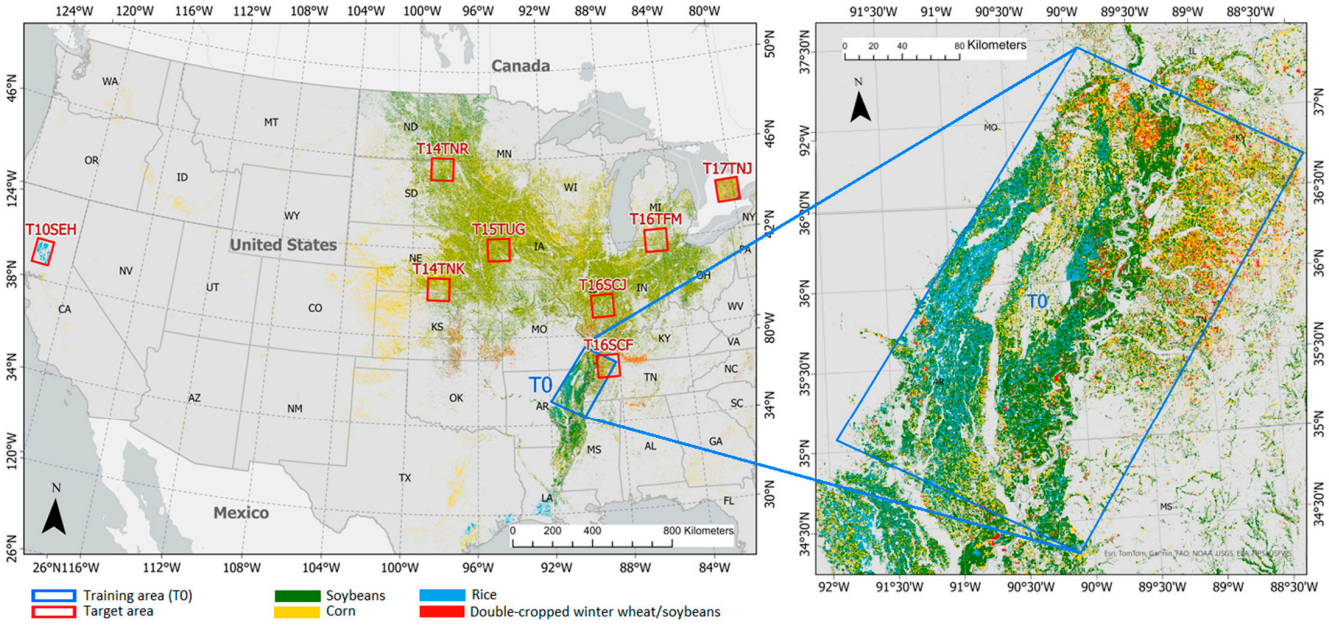


Figure 1. Study area encompassing the designated training area in the Mississippi River Delta (T0) and various target areas across the United States and Canada.

2.2. Remote Sensing Data

Leveraging the expansive temporal reach of Sentinel-2's high-resolution multispectral imagery [5,17,35,36], this research meticulously tracked the phenological stages of the major crops in the training and target area throughout the year 2021. Utilizing the GEE [15,16,37], twelve monthly composite datasets were processed from 2131 tiles, thoroughly covering an area of about 57,596 square kilometers for the training area. This strategic use of Sentinel-2, part of the European Space Agency's Copernicus mission, afforded us detailed and consistent observations. The data's breadth, combined with the satellite's rapid revisiting capability, provided the comprehensive and timely surveillance necessary for precision crop segmentation analysis, capturing the intricate dynamics of crop growth across the agricultural landscape.

For the target areas, Sentinel-2 data in the GEE are provided in tiles, each covering an area of approximately 100 km by 100 km. However, the T10SEH area, which includes California's Sacramento Valley, the second-largest rice-producing state in the US, was covered extensively due to its large agricultural expanse [23,38], resulting in 644 tiles. Other target areas had varying numbers of tiles: T14TNK had 71 tiles, T14TNR had 143 tiles, T15TUG had 72 tiles, T16SCF had 145 tiles, T16SCJ had 71 tiles, T16TFM had 72 tiles, and T17TNJ (Canada) had 69 tiles.

2.3. Ground Truth Data

The 2021 CDL from the United States Department of Agriculture (USDA) was integrated as a critical resource for ground truth verification [13,14]. The CDL provides detailed annual crop cover information across the entire USA, supplying classification maps derived from satellite imagery and agricultural surveys. The accompanying charts in Figures 2 and 3 illustrate the distribution of various land cover categories, including major crops like soybeans, corn, rice, and Dbl-Crop. Additionally, for the T17TNJ (Canada) area, the Annual Crop Inventory (ACI) from Agriculture and Agri-Food Canada (AAFC) [24,39,40] was utilized as another source of ground truth data.

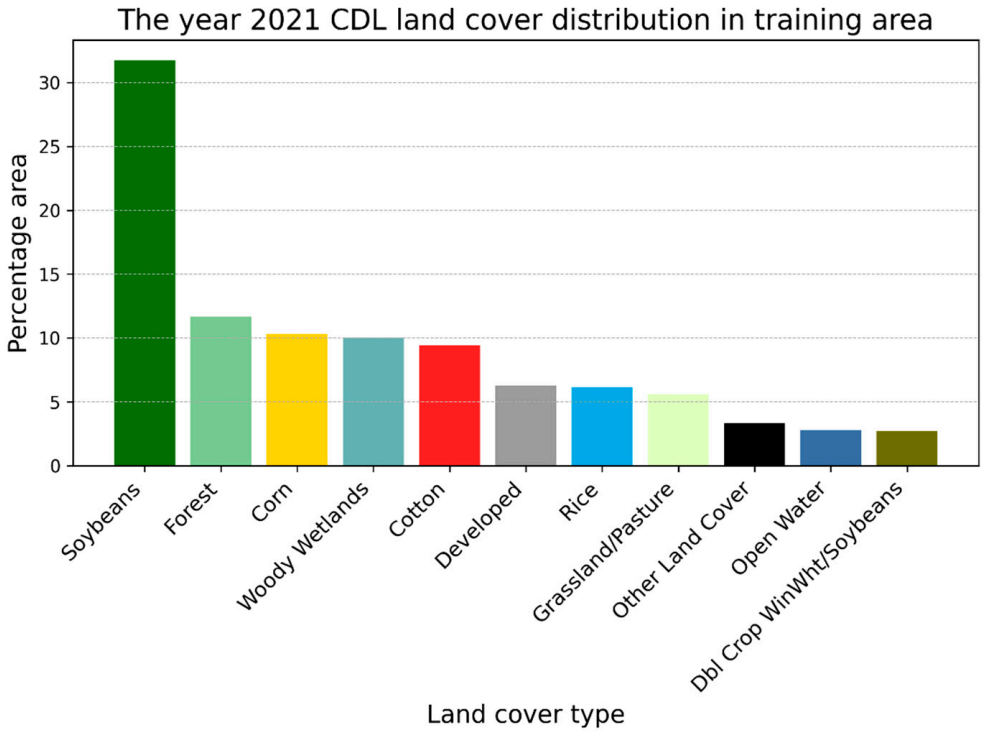


Figure 2. CDL land cover distribution in the training area in 2021, with soybeans as the predominant crop, followed by corn, cotton, rice, and Dbl-Crop.

As shown in Figure 3, for the target area T10SEH, where rice is the major crop, the time series optimization was conducted to align the training data with the phenological stages of rice. For target areas T14TNK, T14TNR, T15TUG, T16SCJ, T16TFM, and T17TNJ (Canada), where corn and soybeans are the major crops, the optimization process was tailored to capture the critical growth periods of these crops. Also, for target areas T17TNJ (Canada) and T16SCF, which have a combination of corn, soybeans, and Dbl-Crop, the time series optimization was performed to reflect the growth patterns of major crops.

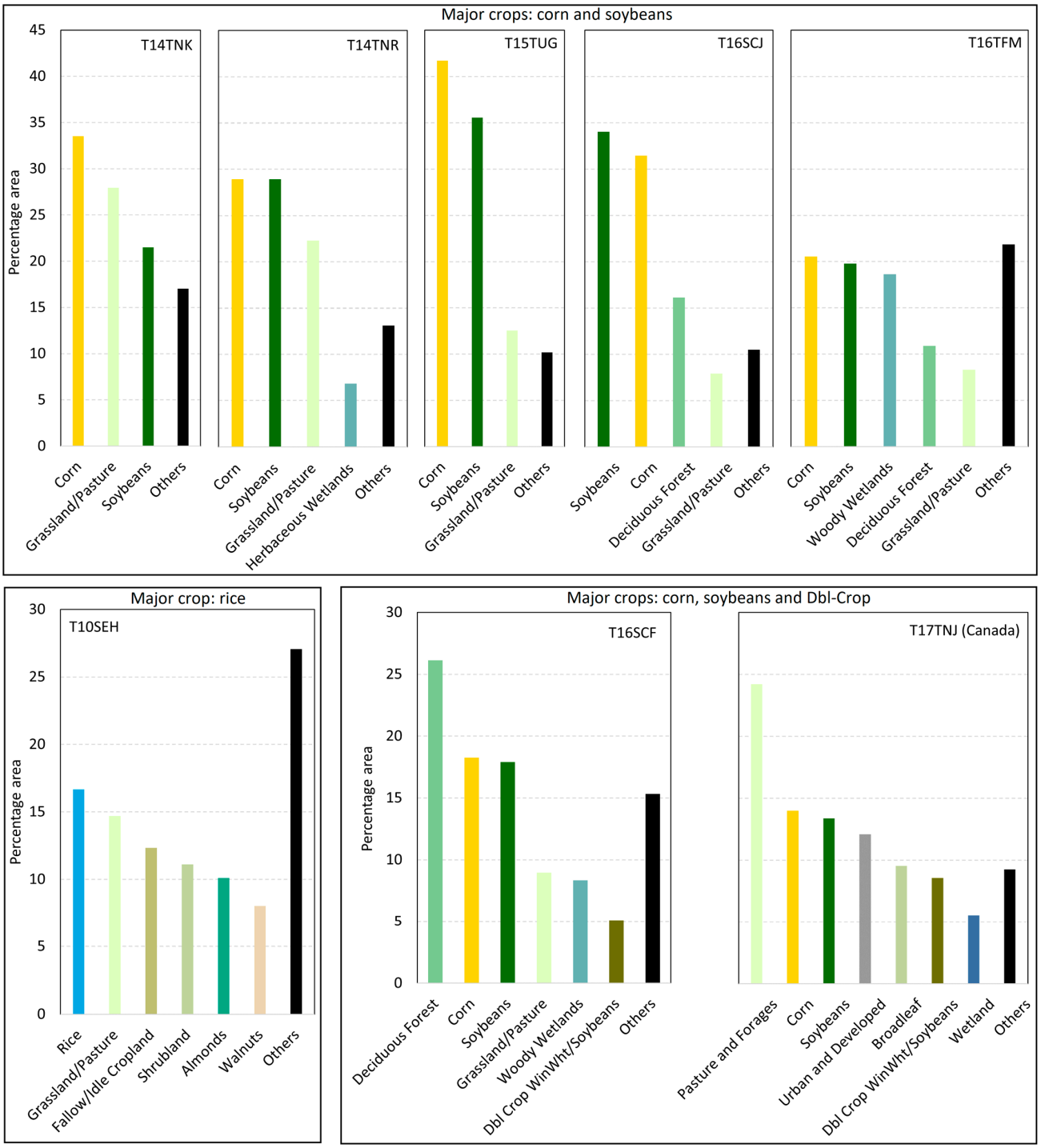


Figure 3. Distribution of land cover categories within each target area based on the 2021 CDL and ACI data, grouped into three categories according to the major crops: corn and soybeans, rice, and a combination of corn, soybeans, and Dbl-Crop.

The NDVI [17,41] was calculated for both the training and target areas using Sentinel-2 imagery processed in the GEE. This index, derived from the red and near-infrared bands of satellite imagery, serves as a critical indicator of vegetation [21,34]. Figure 4 illustrates the NDVI calculations for the training area in 2021, tracking the phenological stages of major crops to support time series alignment for transfer learning in crop classification and segmentation.

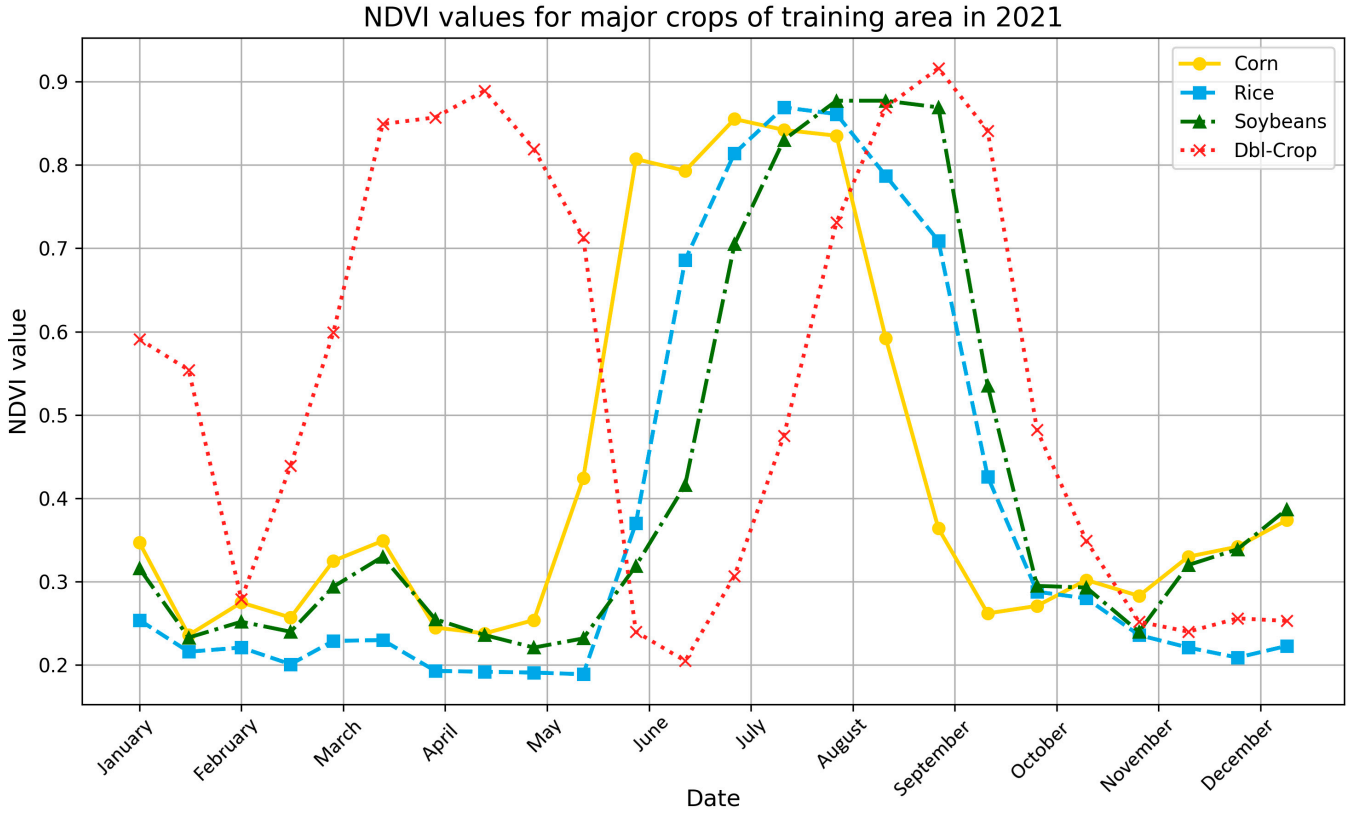


Figure 4. NDVI profiles for major crops such as corn, rice, soybeans, and Dbl-Crop in the training area throughout the year 2021.

2.4. Alignment of Phenological Stages

To address phenological variations across large areas during model transfer, this study introduced the AMM method using NDVI profile matching to synchronize crop growth stages between the training and target areas. By identifying predominant crops in the target region and optimizing time series data in the training area, the method ensured that training data captured critical growth periods of relevant crops. The process of the transfer DL model is illustrated in the flow chart in Figure 5, detailing the sequential steps involved. Initially, Sentinel-2 imagery is acquired for both the training and target areas. The next step involves preprocessing the data using cloud and shadow masking techniques to generate monthly composites. Following this, NDVI profiles are extracted for the training area, and the optimized time series is identified by focusing on the major crops of the target area. These optimized time series are then marked on the NDVI profile of the training area. The NDVI profile of the target area is then extracted and matched to the training area's profile using the AMM method, ensuring the alignment of phenological stages. After this alignment, the corresponding time series of the target area is identified and used for segmentation with the model trained on the optimized time series in the training area. The result is a map of the major crops in the target area, accompanied by the corresponding accuracy metrics to evaluate the segmentation performance. The main steps in the methodology are explained in detail below.

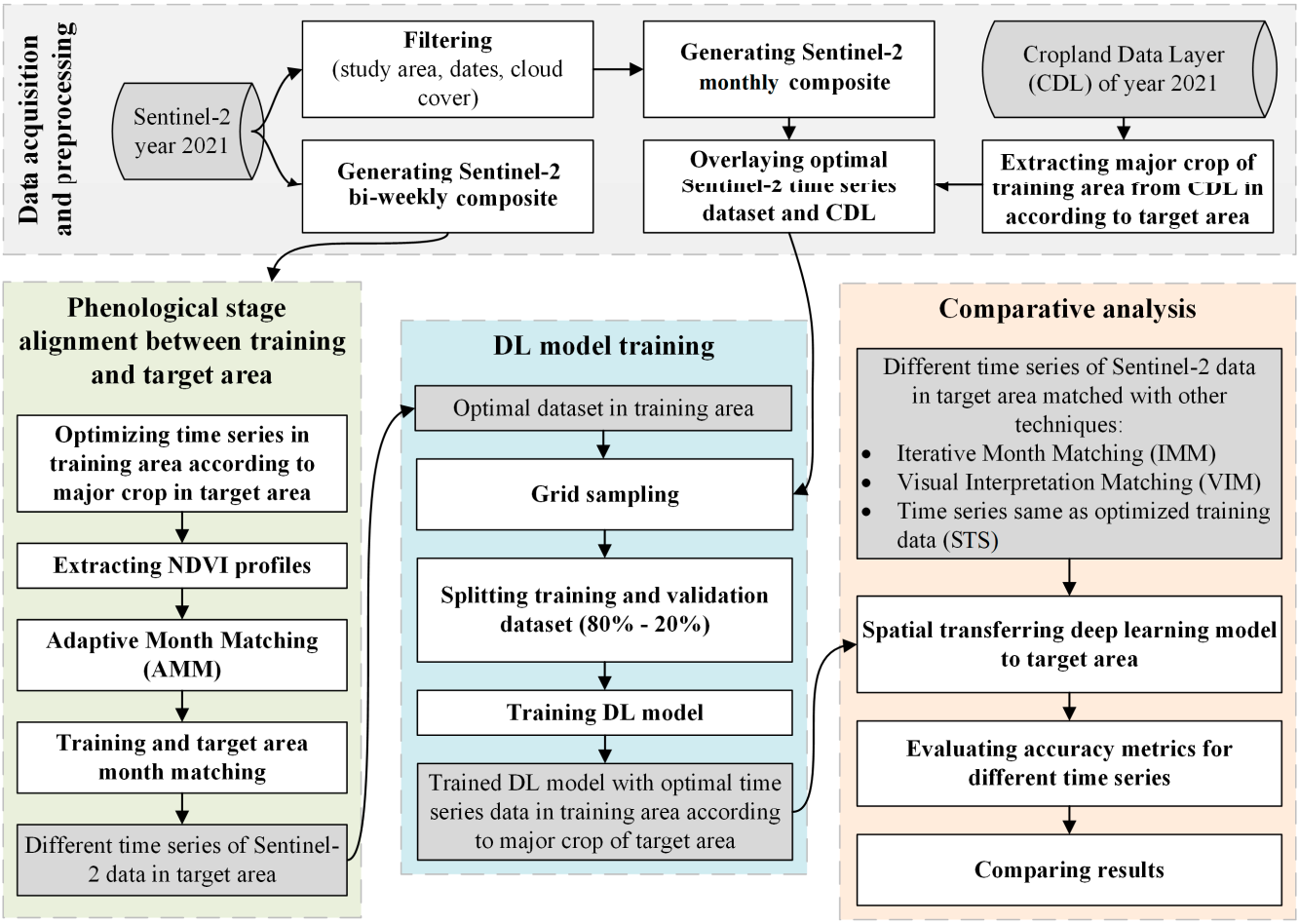


Figure 5. The flowchart illustrates the process of data acquisition and preprocessing, alignment of phenological stages, DL model training, and a comparative analysis of different methods for an effective transfer learning model in crop segmentation.

AMM is a novel method designed to align time series data between two areas based on their NDVI changes, enhancing the transfer of trained DL models for crop segmentation. By ensuring that the phenological time series of the training and target areas are optimally matched, AMM is designed to improve model performance in the target area. A flowchart in Figure 6 visually describes the key steps of the AMM method, which include the following:

- Optimizing time series in training area: Use an iterative Multi-Criteria Decision Analysis (MCDA) [42] to optimize the time series data in the training area based on the major crops in the target area.
- Extracting NDVI profiles: Compute and average the NDVI profiles from bi-weekly Sentinel-2 time series data for each major crop in both the training and target areas.
- Selecting candidate months: Calculate NDVI changes for candidate months in the target area, including the same months as the training time series, as well as one month before and one month after.
- Compute differences and slopes: Calculate the absolute differences and slope differences between each month in the optimal training time series and the corresponding candidate months in the target area.
- Calculate combined score: Normalize the differences and slopes and compute combined scores.
- Identify best match: Select the candidate month with the lowest combined score as the best match.

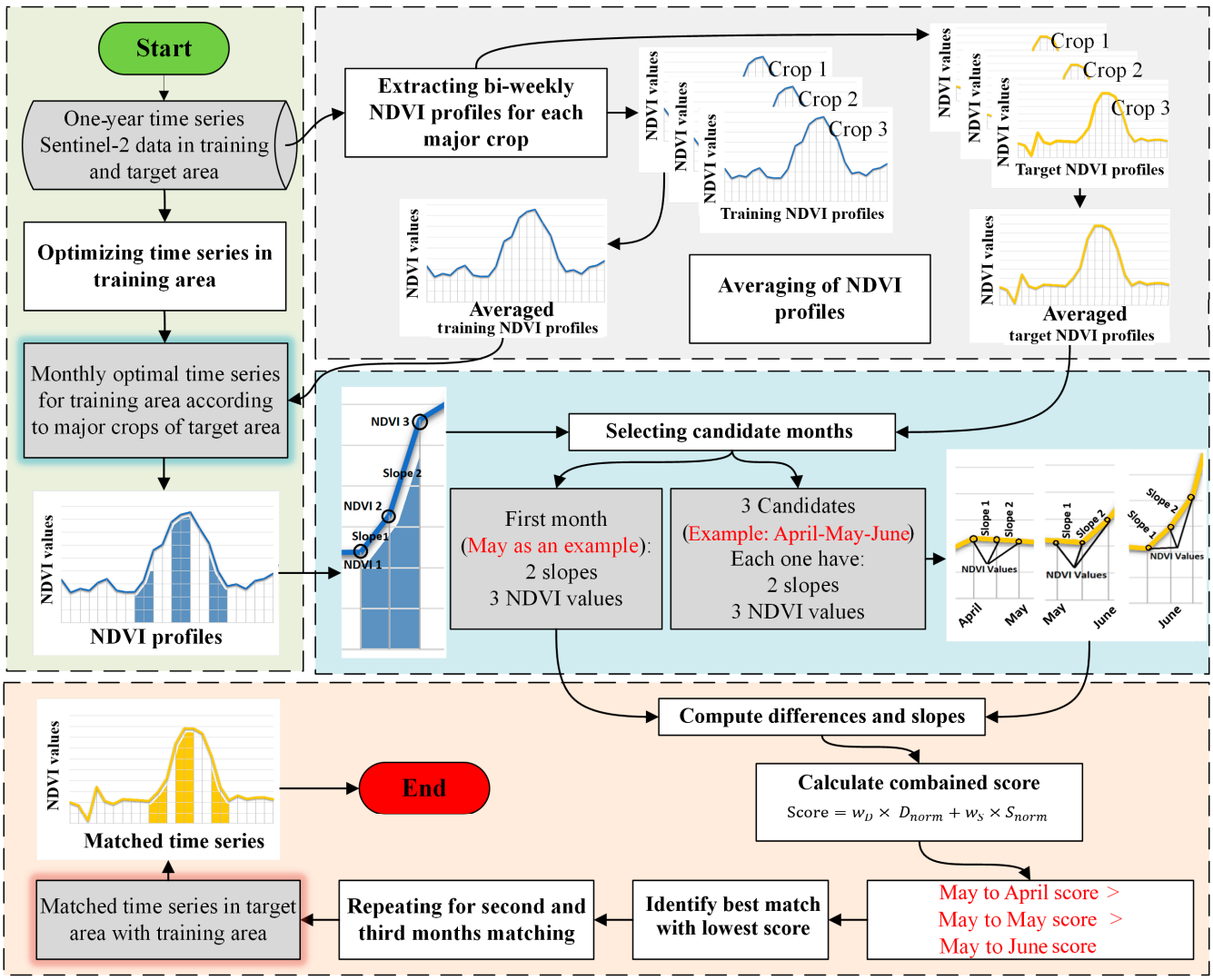


Figure 6. Flowchart of the AMM method illustrating the process of extracting and averaging NDVI profiles for major crops and determining the best time series in the target area matched to the training area time series. The blue color represents the training NDVI profiles, while the yellow color represents the target NDVI profiles. In the flowchart, the months of May, April, and June are used as examples.

The optimal time series in the training area was determined using iterative MCDA based on the major crops of the target area. This optimization ensured the segmentation model accurately captured the phenological stages of these crops. The process began by identifying the primary crops in the target area through agricultural reports. Each time series was evaluated based on accuracy metrics of major crops. Initially, core months critical for capturing these stages were identified using MCDA, which evaluated DL segmentation accuracy from models trained with different monthly Sentinel-2 composites. Accuracy metrics were normalized to ensure comparability across different scales and were assigned equal weight. The MCDA process consolidated accuracy metrics, calculating the relative closeness of each time series option to an ideal solution [34]. This approach ranked and selected the optimal time series based on UA, PA, and F1-score. The distance to the best ideal (D^+) and worst ideal (D^-) was calculated for each alternative and the relative closeness (RC) was determined to rank alternatives. Subsequent iterations expanded the core time series by incorporating additional months and re-evaluating the segmentation

accuracy and *RC*. This iterative process continued until the optimal time series in the training area was identified. The *RC* for an alternative (*a*) was calculated as

$$RC_a = \frac{D_a^-}{D_a^+ + D_a^-} . \tag{3}$$

The criterion for halting the iterative process is the attainment of a time series where the incremental accuracy gain does not justify the additional complexity and data processing demand. The stopping criterion for the iterative exploration process can be represented as

$$\Delta RC_{t+1} = RC_{W_t \cup W_{t+1}} - RC_{W_t}, \tag{4}$$

$$\Delta RC_{t+1} < \varepsilon, \tag{5}$$

where RC_{W_t} is the *RC* of the current time series W_t, which includes t months. $RC_{W_t \cup W_{t+1}}$ represents the *RC* of the expanded time series after adding the next month and ΔRC_{t+1} is the incremental gain in the *RC* when expanding from W_t to W_{t+1} by including the next month. If ΔRC_{t+1} is less than the predefined threshold ε, the iterative process is terminated. ε is the minimum acceptable incremental gain in accuracy to justify the additional complexity of adding another month, set at 1%.

After extracting the average NDVI profiles, the NDVI changes over three consecutive periods (spanning parts of adjacent months) were calculated for each selected month. These changes were derived by taking the difference between NDVI values at specified intervals, providing a sequence of NDVI change rates that characterize the vegetation dynamics during that month. Mathematically, the change rate for month *M* was represented as

$$\Delta NDVI_M = [NDVI_{M+15} - NDVI_M, NDVI_{M+30} - NDVI_{M+15}, NDVI_{M+45} - NDVI_{M+30}] . \tag{6}$$

Next, the method evaluated potential corresponding months in the target dataset by calculating similar sequences of NDVI changes. For each candidate test month, the absolute differences and slopes between its NDVI change rates and those of the training month were computed. The absolute difference *D* and slope difference *S* were calculated as

$$D = \sum_{i=1}^{3} |\Delta NDVI_{Train}[i] - \Delta NDVI_{Test}[i]|, \tag{7}$$

$$S = \sum_{i=1}^{2} | \, (\Delta NDVI_{Train}[i+1] - \Delta NDVI_{Train}[i]) - (\Delta NDVI_{Test}[i+1] - \Delta NDVI_{Test}[i]). \, | \tag{8}$$

These differences and slopes were then normalized to ensure comparability:

$$D_{norm} = \frac{D - D_{min}}{D_{max} - D_{min}} , \tag{9}$$

$$S_{norm} = \frac{S - S_{min}}{S_{max} - S_{min}} . \tag{10}$$

By combining these normalized values with weighted sums, the AMM method scored each candidate month, ultimately selecting the month with the lowest combined score as the best match, where w_D and w_S are the weights assigned to the normalized difference and slope, respectively:

$$\text{Score} = w_D \times D_{norm} + w_S \times S_{norm}. \tag{11}$$

For example, to align May from the training region to the target region, the AMM method first calculates the NDVI changes for May in the training dataset. It then compares potential matches in the target region, such as April, May, and June. The method computes the absolute differences and slope differences between the NDVI changes for May in the training region and April, May, and June in the target region. These values are then normalized and combined to calculate a score for each candidate month. The month with the lowest combined score, say June in this case, would be identified as the best match for May in the training area. This process is repeated for the other months in the training time series to find the best match for each. This ensures that the training model is aligned with the phenological stages in the target region for optimal performance.

2.5. Data Preparation

To prepare data for training the DL model, a comprehensive preprocessing workflow was implemented on Sentinel-2 imagery. This involved cloud and shadow masking using Sentinel-2's cloud probability data and scene classification layer [34,43]. To ensure the imagery's suitability for detailed crop segmentation, the Sentinel-2 bands B01, B09, and B10, corresponding to coastal aerosols, water vapor, and cirrus clouds, were excluded due to their 60 m resolution, which is insufficient for accurately capturing small surface features. As a result, each monthly composite used for training the DL model consists of 10 bands [34]. The monthly composite images were generated, with all preprocessing steps applied in the GEE environment to ensure data quality and reduce cloud interference. Concurrently, the CDL was refined to focus on major crops in both the target and training areas, masking out other crops and adjusting the CDL to align with the spatial coverage of the Sentinel-2 imagery composite. Data preparation included overlaying Sentinel-2 composite captures onto the CDL and using grid sampling to generate 256 × 256-pixel samples. Less informative samples were filtered out, and the refined sample set was split into 80% training and 20% validation groups [34].

2.6. DL Model Training

For segmenting crops from Sentinel-2 composite data, we employed the PyTorch semantic segmentation U-Net [41] model with a ResNet-34 encoder [23,34]. The U-Net architecture is renowned for its effectiveness in semantic segmentation across diverse applications [44], while the ResNet-34 encoder provides a strong foundation for feature extraction [45,46]. Although the encoder weights were initialized using ImageNet, this primarily served the purpose of providing a starting point for training, rather than relying on random initialization. The architecture included a standard U-Net decoder with batch normalization, adapted to the specific requirements of our dataset. The hyperparameter tuning process involved experimenting with different learning rates, batch sizes, and other model parameters. The Adam optimizer was employed with an initial learning rate of 0.01, determined through a hyperparameter tuning process. The models were trained using cross-entropy loss and incorporated early stopping with a patience of 50 epochs to prevent overfitting. The training process was capped at 500 epochs; however, in our implementation, all training converged within 400 epochs, with most completing near 300 epochs, demonstrating that this limit was sufficient.

The input channels were aligned with the Sentinel-2 bands and the number of months in the time series. Since each monthly composite consists of 10 bands, the time series included one, two, and three months, resulting in 10, 20, and 30 input channels, respectively. This corresponds to the number of months multiplied by the 10 bands per month. For each major crop group in the target area, five different models were trained. This included three models for the three best individually matched months, one model for the best two-month

matched time series, and another for the entire three-month time series. The major crops in the target area were categorized into three groups, as shown in Figure 3. Consequently, 15 models were trained across five different data time series for the three major crop groups.

2.7. Evaluation

To evaluate the performance of the crop segmentation DL models, a suite of accuracy metrics was employed: Overall Accuracy (OA), User's Accuracy (UA), Producer's Accuracy (PA), and F1-score (F1) [2,41,47]. OA, UA, and PA are presented as percentages, while the F1-score ranges between 0 and 1.

$$\mathrm{OA} = \frac{TP + TN}{n}, \tag{12}$$

$$\mathrm{UA} = \frac{TP}{TP + FP}, \tag{13}$$

$$\mathrm{PA} = \frac{TP}{TP + FN}, \tag{14}$$

$$\mathrm{F1} = 2 \times \frac{\mathrm{UA} \times \mathrm{PA}}{\mathrm{UA} + \mathrm{PA}}, \tag{15}$$

where n is the total number of instances, TP is the correctly predicted positive instances, and TN is the correctly predicted negative instances [41]. FP represents the incorrectly predicted positive instances, and FN stands for the incorrectly predicted negative instances.

To ensure a comprehensive comparison, we used three different month-matching techniques: Iterative Month Matching (IMM), Visual Interpretation Matching (VIM), and matching the same time series months in the target area as in the training area (STS). The IMM process, using MCDA, identified the optimal time series for the major crops in the training area by training multiple models with different combinations of optimal months. These models were applied to the target area, and their performance was tested with different monthly data inputs. The optimal time series for the target area was identified through an iterative process, with accuracy evaluated by MCDA. IMM serves as a benchmark to assess AMM's performance but is less practical for transfer learning due to its reliance on detailed cropland maps, which are often unavailable. In contrast, AMM offers a more adaptable approach, eliminating the need for specific local crop information, making it a practical solution for transfer learning across diverse regions.

In addition to automated methods, VIM provided a visual alignment of NDVI profiles between the training and target areas through expert interpretation. Agronomists and remote sensing specialists visually matched the phenological stages, helping identify the optimal time series for the target area. The VIM results were then compared with those of automated methods, offering a well-rounded evaluation of the techniques. Additionally, we compared our results with other studies. This multi-faceted approach allowed us to benchmark and evaluate the effectiveness and robustness of AMM.

3. Results

In this section, we present the results of time series matching for the target areas' Sentinel-2 monthly composite data and the training areas' Sentinel-2 data. The target areas were divided into three groups based on their predominant crops: a corn and soybeans group, a corn, soybeans, and Dbl-Crop group, and a rice group. For each crop group, time series data were optimized in the training area for one-, two-, and three-month time series. Different DL models were trained with various time series for different classes. This approach is designed to ensure that the training data effectively represent the critical

growth stages of these crops. The following sections present the results of transferring these DL models to different target areas.

3.1. Transfer DL Model for Corn and Soybeans

For the first group of crops, corn and soybeans, the optimization process identified the optimal time series combinations for segmentation accuracy in the training area. Initially, July was found to be the optimal core month. Adding September formed the best two-month combination, and including May resulted in the optimal three-month time series of May, July, and September. As a result, we trained five DL models for corn and soybeans segmentation. Three models were trained using Sentinel-2 composite data from May, July, and September individually. Another model was trained using the two-month combination of July–September, and the final model was trained using the optimal three-month combination of May, July, and September.

We also transferred these five DL models to T17TNJ (Canada) to evaluate their performance without considering Dbl-Crop, as the results for Dbl-Crop in this target area were unsatisfactory. Specifically, Dbl-Crop had the lowest F1-score, nearly zero, indicating that the transferred DL model was unable to properly classify Dbl-Crop, which negatively impacted the overall training process in T17TNJ (Canada). Additionally, analysis revealed that the NDVI profile of winter wheat in the ACI from AAFC in T17TNJ (Canada) closely resembles that of Dbl-Crop in the CDL, despite the ACI not listing Dbl-Crop as a crop category. This discrepancy further contributed to the poor performance for Dbl-Crop in this target area.

Table 1 presents the detailed results of transferring the DL models to the target areas where the major crops are corn and soybeans. For the AMM method, the optimal matches were determined to be June–July–October for T14TNK, June–July–September for T14TNR, T16SCJ, T16TFM, and T17TNJ (Canada), and June–August–October for T15TUG, according to the absolute differences and slope differences. For the IMM technique, in the T14TNK, T14TNR, and T16TFM target areas, the optimal matches were June–July–September. In T15TUG, IMM identified June–August–October as the optimal match, and for T16SCJ, the optimal match of Sentinel-2 composite images was June–August–September. In T17TNJ (Canada), the optimal match was June–August–October.

Table 1. Results of transfer DL model trained on optimal three months' time series (May–July–September) for the training area for corn and soybeans segmentation in target areas including T14TNK, T14TNR, T15TUG, T16SCJ, T16TFM, and T17TNJ (Canada). In table: Overall Accuracy (OA), Producer's Accuracy (PA), User's Accuracy (UA), F1-score (F1), Adaptive Month Matching (AMM), Iterative Month Matching (IMM), Visual Interpretation Matching (VIM), and same time series as training area (STS). These methods are explained in more detail in Section 2.

Target Areas	Matching Method	Months Matched to May–July–September	OA (%)	Corn [1]			Soybeans			Other		
				UA (%)	PA (%)	F1	UA (%)	PA (%)	F1	UA (%)	PA (%)	F1
T14TNK	STS	May–July–September	60.2	**98.1** [2]	16.0	0.27	61.2	63.2	0.58	59.4	89.3	0.69
	IMM, VIM	June–July–September	69.4	97.1	34.8	0.50	**68.6**	69.8	**0.66**	65.5	**92.0**	0.74
	AMM	June–July–October	**72.9**	94.8	**49.6**	**0.63**	57.2	**78.1**	0.65	**73.1**	82.6	**0.77**
T14TNR	STS	May–July–September	57.6	**96.8**	33.7	0.48	47.8	81.9	0.58	60.5	56.4	0.56
	VIM	June–July–October	**65.6**	74.1	45.2	0.55	**87.4**	69.8	**0.77**	48.6	**88.3**	**0.62**
	IMM, AMM	June–July–September	64.8	95.7	**52.1**	**0.67**	53.3	**86.8**	0.64	**66.7**	58.6	**0.62**
T15TUG	STS	May–July–September	61.4	**98.7**	32.0	0.48	86.3	66.3	0.74	41.1	92.4	0.56
	VIM	June–August–September	61.7	97.8	37.3	0.52	**92.8**	59.5	0.71	41.3	**94.0**	0.55
	IMM, AMM	June–August–October	**80.1**	83.4	**92.2**	**0.87**	81.3	**71.9**	**0.76**	**73.1**	66.2	**0.69**
T16SCJ	VIM, STS	May–July–September	61.9	**98.6**	10.4	0.18	**92.5**	66.1	0.76	50.9	**98.8**	0.66
	IMM	June–August–September	**80.4**	97.1	**57.4**	**0.71**	83.8	**83.6**	**0.83**	**71.2**	94.8	**0.80**
	AMM	June–July–September	75.9	98.1	48.9	0.64	87.0	77.4	0.81	63.3	95.3	0.75

Table 1. *Cont.*

Target Areas	Matching Method	Months Matched to May–July–September	OA (%)	Corn [1] UA (%)	Corn [1] PA (%)	Corn [1] F1	Soybeans UA (%)	Soybeans PA (%)	Soybeans F1	Other UA (%)	Other PA (%)	Other F1
T16TFM	STS	May–July–September	71.3	**96.7**	6.3	0.11	**80.8**	48.7	0.60	69.5	**99.0**	0.81
	AMM, IMM, VIM	June–July–September	**76.9**	95.7	**27.5**	**0.42**	79.5	**59.5**	**0.67**	**74.6**	97.9	**0.85**
T17TNJ (Canada)	STS	May–July–September	77.0	**97.1**	5.5	0.09	68.8	23.4	0.34	76.5	**99.6**	0.86
	IMM	June–August–October	**83.9**	95.9	**26.9**	**0.39**	**78.3**	**62.5**	**0.68**	**83.9**	98.4	**0.90**
	VIM	June–July–October	77.1	95.0	11.9	0.19	72.3	31.8	0.42	77.5	99.3	0.86
	AMM	June–July–September	76.9	95.6	8.5	0.16	67.2	28.8	0.39	77.1	99.2	0.87

[1] Results for different crops are separated with distinct backgrounds. [2] The bolded accuracy metric values indicate the best results in each separated row.

The NDVI chart in Figure 7 compares the bi-weekly NDVI profiles of the training and target areas for corn and soybeans in 2021. These profiles were used for visual interpretation to match (VIM) the corresponding months between the training and target areas. Based on the visual interpretation (VIM) of NDVI profiles between the training and target areas, we concluded that the optimal matched time series for the target area corresponding to the May–July–September series in the training area was June–July–September for T14TNK and T16TFM, June–August–September for T15TUG, and June–July–October for T14TNR and T17TNJ (Canada). For the T16SCJ target area, May–July–September, the same time series as the training area, was identified as the optimal visual match.

In target area T14TNK, the AMM method identified June–July–October as the optimal time series, achieving an overall accuracy of 73% and demonstrating superior performance in corn segmentation, with a PA of 95% and UA of 50%. This combination outperformed the May–July–September time series optimized in the training area, which had a lower OA of 60%. In T14TNR, the June–July–September combination, determined by the AMM and IMM methods, achieved an OA of 64.8%, compared to an OA of 57.65% achieved with the STS method, highlighting the importance of tailoring the time series to the target area's specific phenological stages. For T15TUG, the highest OA was achieved with the June–August–October combination as determined by the AMM and IMM methods, showing significant improvement in corn and soybeans segmentation, with approximately 19% higher accuracy compared to the STS method. The AMM method again provided consistently high-performing results. In T16SCJ, both the IMM and AMM methods demonstrated high accuracy, with IMM performing best for corn segmentation in August and AMM excelling in soybeans segmentation (F1-score of 0.77 for June–July–September). For T16TFM, the June–July–September period demonstrated superior performance across metrics, with an approximately 5% improvement compared to the STS method, confirming the robustness of the AMM method.

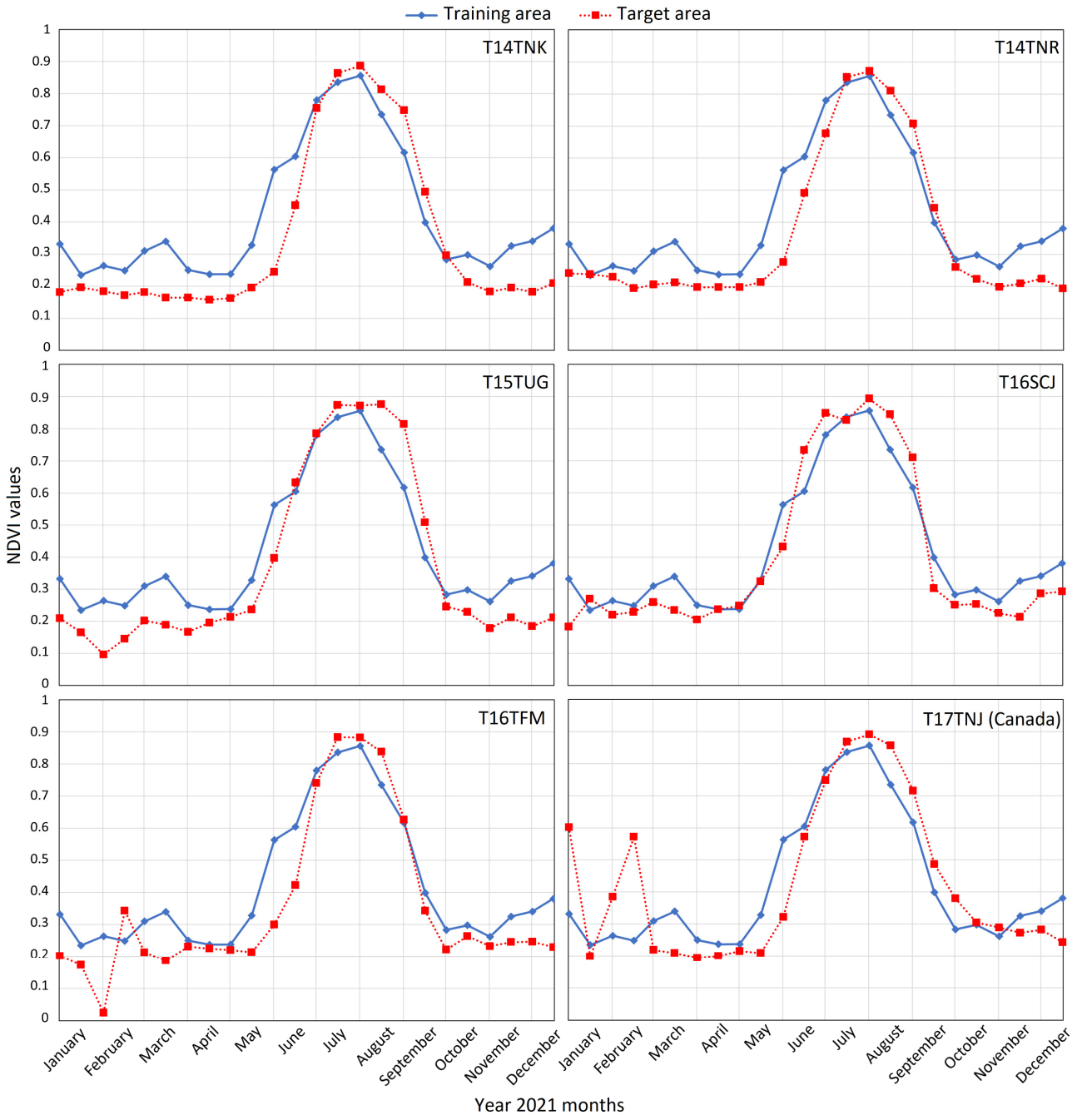


Figure 7. NDVI profiles of the training area (blue) and target area (red) for corn and soybeans in 2021. The chart illustrates the bi-weekly NDVI values for each area.

In T17TNJ (Canada), the IMM method achieved the best overall performance for the June–August–October time series, with an OA of 83.9%, while the AMM method identified June–July–September as the optimal time series, delivering a moderate performance with an OA of 76.9%, which is 7% lower than that achieved using the IMM method. Overall, AMM's adaptability to specific phenological stages in the target areas resulted in an enhanced segmentation accuracy, showcasing its effectiveness in improving transfer learning for crop segmentation. The AMM method consistently demonstrated strong performance in the US target areas, achieving improvements ranging from approximately 5% to 19% compared to the STS method, underscoring its reliability for time series matching in diverse regions.

Figure 8 illustrates the comparison of DL model performance for different time series combinations in various target areas. By visually comparing the results of the DL models

with the CDL, it becomes evident that, in some cases, increasing the time series from one month to three months led to a decline in segmentation quality. This effect is particularly noticeable in the T16SCJ, T16TFM, and T17TNJ (Canada) regions. However, with a single-month time series, the boundaries of croplands tended to blend with adjacent roads. In the T14TNK region, adding October composite images caused much of the area to be incorrectly segmented as soybeans, suggesting that October may not have been an appropriate choice. Conversely, in T14TNR, adding July and September improved the segmentation quality. For T15TUG, August was a key month, with the addition of other months leading to slight improvements in segmentation performance based on visual comparisons.

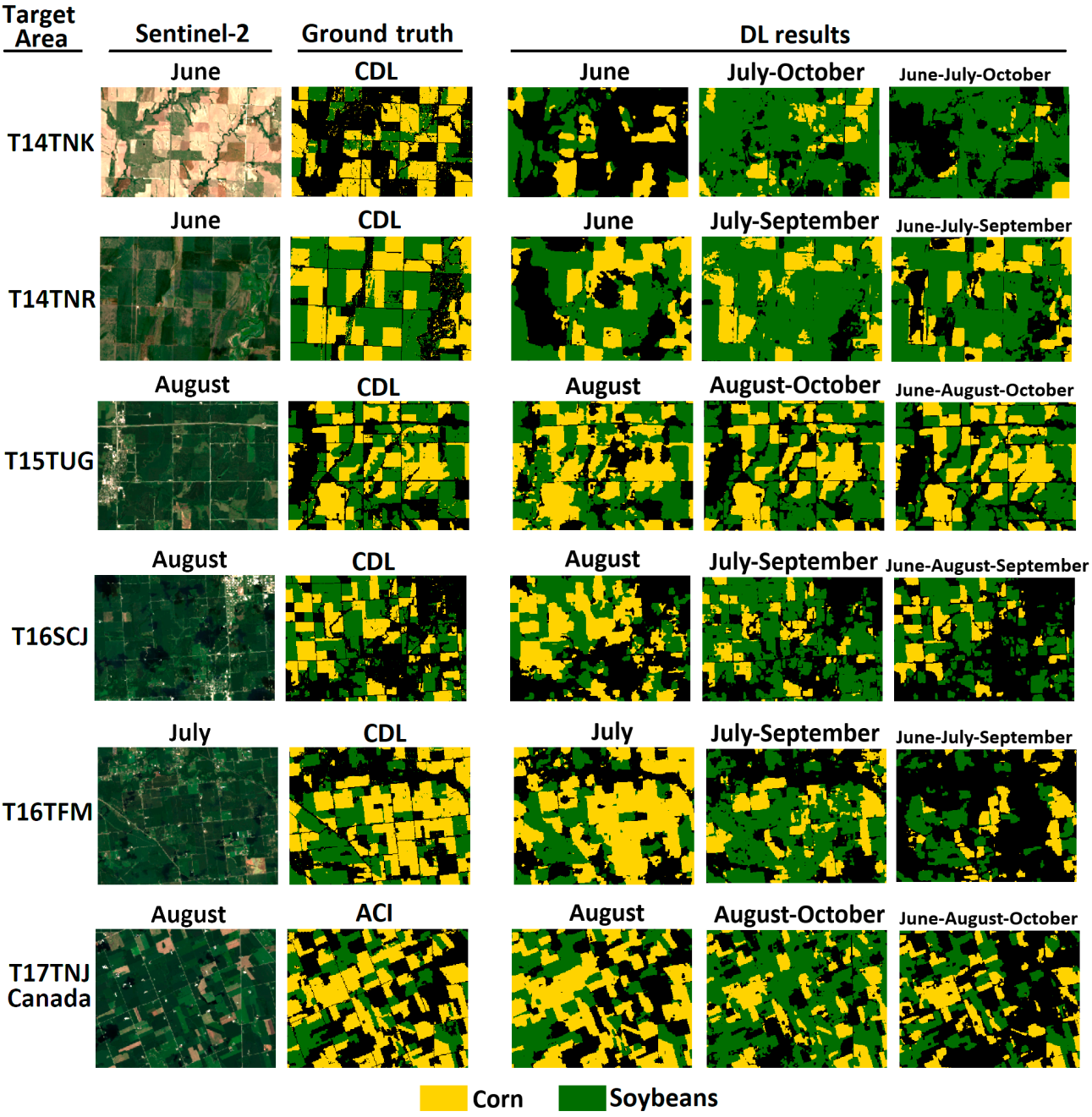


Figure 8. Visual comparison of results of transfer DL models for corn and soybeans segmentation in the T14TNK, T14TNR, T15TUG, T16SCJ, T16TFM, and T17TNJ (Canada) target areas showing Sentinel-2 composite images, ground truth data (CDL and ACI), and DL results for one-, two-, and three-month time series matching.

3.2. Transfer DL Model for Corn, Soybeans, and Dbl-Crop

For the second group of crops, which includes corn, soybeans, and Dbl-Crop, the optimization process used the MCDA approach to identify the optimal time series combinations. June was identified as the optimal first month. After adding subsequent months, the optimal combinations were June–September and April–June–September. The trained models for each time series combination were saved for transfer to the target area, resulting in five distinct DL segmentation models.

In T16SCF, which has a significant overlap with the training area, the NDVI profiles for both areas are notably similar, as depicted in Figure 9. Given this overlap, the results of various time series matching methods align closely with those for the training area, identifying April–June–September as the optimal time series for these crops. The transferred DL model results in target area T16SCF are summarized in Table 2. The IMM, VIM, and AMM methods all converged on April–June–September as the optimal time series, achieving a 90% OA.

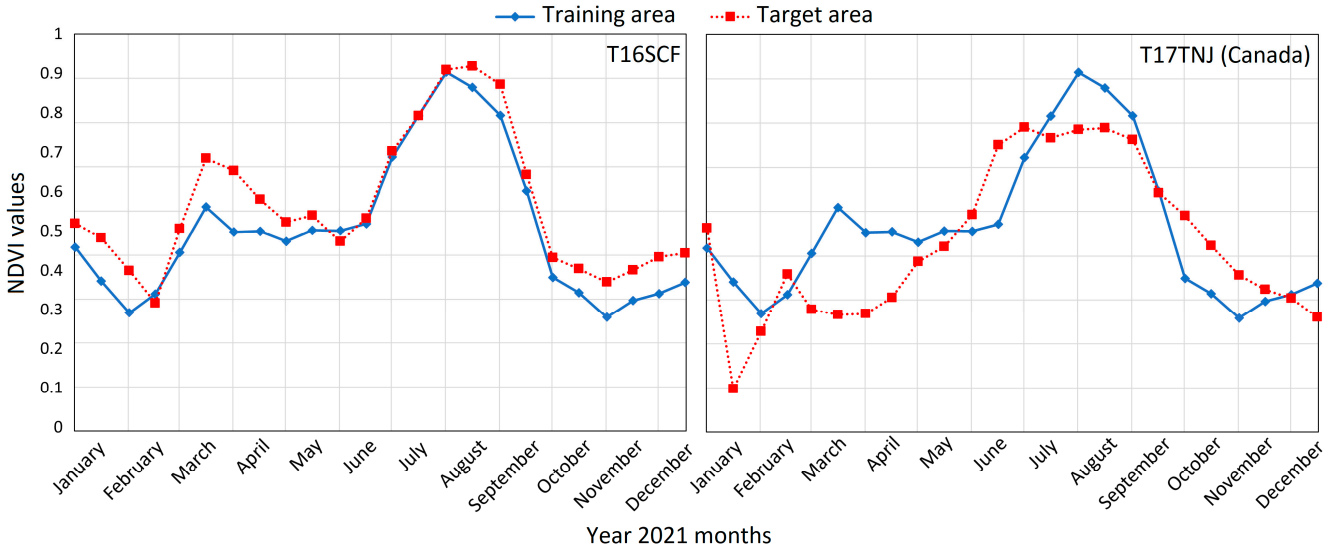


Figure 9. NDVI profiles of the training area (blue) and target areas (red) T16SCF and T17TNJ (Canada) for corn, soybeans, and Dbl-Crop.

Table 2. Result of transfer DL models for corn, soybeans, and Dbl-Crop segmentation in target areas T16SCF and T17TNJ (Canada). The IMM, AMM, VIM, and STS methods are explained in more detail in Section 2.

Target Areas	Matching Method	Months Matched to April–June–September	OA (%)	Corn [1] UA (%)	Corn PA (%)	Corn F1	Soybeans UA (%)	Soybeans PA (%)	Soybeans F1	Dbl-Crop UA (%)	Dbl-Crop PA (%)	Dbl-Crop F1	Other UA (%)	Other PA (%)	Other F1
T16SCF	IMM, AMM, VIM, STS	April–June–September	90.0	93.8	76.0	0.84	82.7	81.5	0.81	79.0	94.7	0.86	91.7	95.2	0.93
T17TNJ (Canada)	STS	April–June–September	63.3	74.5	**90.1** [2]	**0.81**	9.2	0.6	0	26.5	**42.0**	**0.31**	4.2	0.7	0.00
	IMM	May–July–August	**72.0**	**86.7**	2.1	0.03	**75.4**	**67.7**	**0.70**	**37.2**	14.7	0.19	**71.4**	**95.1**	**0.81**
	VIM	March–May–October	50.8	39.0	8.3	0.13	14.2	28.0	0.18	0.5	0.7	0.00	64.1	72.8	0.67
	AMM	March–June–October	51.9	35.8	32.1	0.32	16.2	29.8	0.20	1.8	0.6	0.00	67.0	66.8	0.66

[1] Results for different crops are separated with distinct backgrounds. [2] The bolded accuracy metric values indicate the best results in each separated row.

In T17TNJ (Canada), the major crops are the same as in T16SCF, but the NDVI profile differs due to geographic distance (see Figure 9). The results of the transferred DL models in this area show that the AMM method achieved a 51.9% OA, performing worse than IMM (72% OA) and STS (63.3% OA), while the VIM method also delivered suboptimal

results, with a 50.8% OA. By analyzing each crop individually, it was found that the NDVI profile of winter wheat in the ACI of AAFC is similar to Dbl-Crop in the CDL, even though the ACI does not list Dbl-Crop. The IMM method was the most effective in this region, achieving an approximately 9% higher OA than the STS method. Its highest segmentation performance was obtained with a model trained on June data from the training area and tested on July data from the target area, reaching a 72.0% OA.

Figure 10 illustrates a comparison of DL model performance for different time series combinations for corn, soybeans, and Dbl-Crop segmentation in T16SCF and T17TNJ (Canada). From a visual comparison of the transferred DL models' results, it can be observed that in the T17TNJ (Canada) target area, increasing the time series from one to two or three months resulted in a significant decline in segmentation quality. The best performance for this area was achieved using the July composite images; however, the addition of September, May, and August did not enhance the segmentation results. In contrast, for the T16SCF area, the segmentation quality improved with the inclusion of additional months, particularly for cropland boundaries, where roads around agricultural fields were misclassified as crops in the single-month segmentation. The difference in performance between these two areas can be explained by their geographical proximity to the training area. T16SCF is much closer to the training area, whereas T17TNJ (Canada) is one of the most distant. Furthermore, the ground truth data for T17TNJ (Canada) were derived from the ACI dataset, rather than the CDL, which may have contributed to the differences in segmentation accuracy.

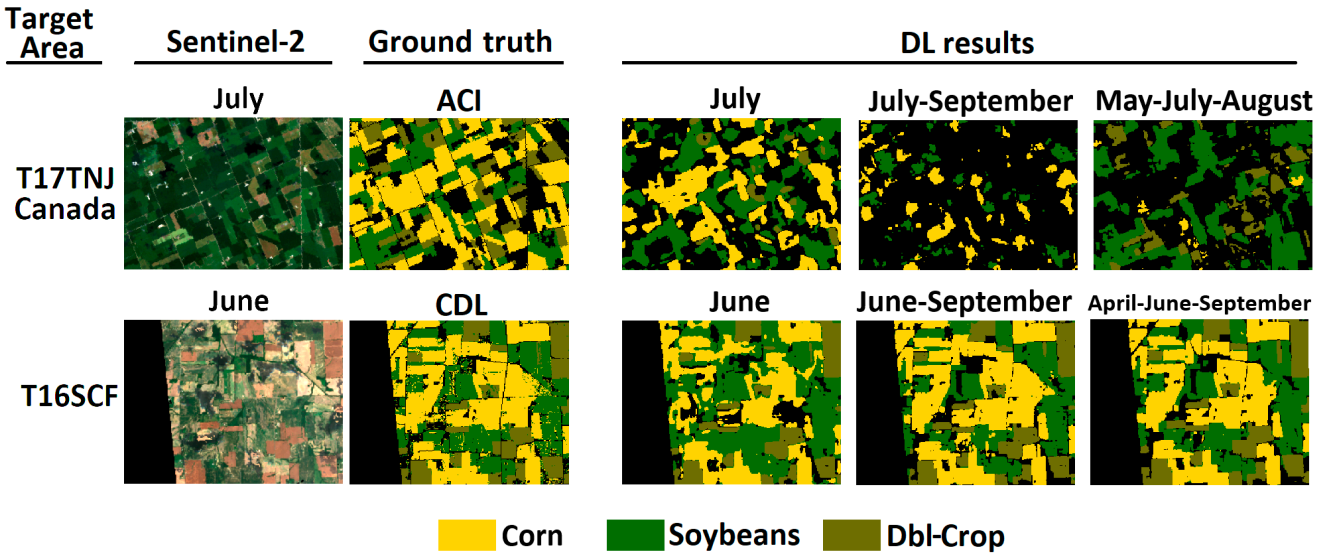


Figure 10. Visual comparison of results of transfer DL models for corn, soybeans and Dbl-Crop in T16SCF and T17TNJ (Canada).

3.3. Transfer DL Model for Rice

For the third group of crops, which focuses on rice, July was identified as the optimal first month. The optimal combination time series were July–September and July–August–September. The trained models for each time series combination were stored for subsequent transfer DL models, yielding five distinct DL segmentation models. For target area T10SEH, which primarily focuses on rice, the time series matching results from different methods are presented in Table 3 and the NDVI profiles of the training and target areas are illustrated in Figure 11.

Table 3. Results of transfer DL models for rice segmentation in the T10SEH target area. The IMM, AMM, VIM, and STS methods are explained in more detail in Section 2.

Target Areas	Matching Method	Months Matched to July–August–September	OA (%)	Rice [1] UA (%)	Rice [1] PA (%)	Rice [1] F1	Other UA (%)	Other PA (%)	Other F1
T10SEH	IMM, STS	July–August–September	97.17	82.58	**93.14** [2]	0.860	98.93	97.54	0.979
	AMM, VIM	June–August–September	**97.88**	**85.36**	92.83	**0.873**	**98.96**	**98.47**	**0.982**

[1] Results of rice and other are separated with distinct backgrounds. [2] The bolded accuracy metric values indicate the best results.

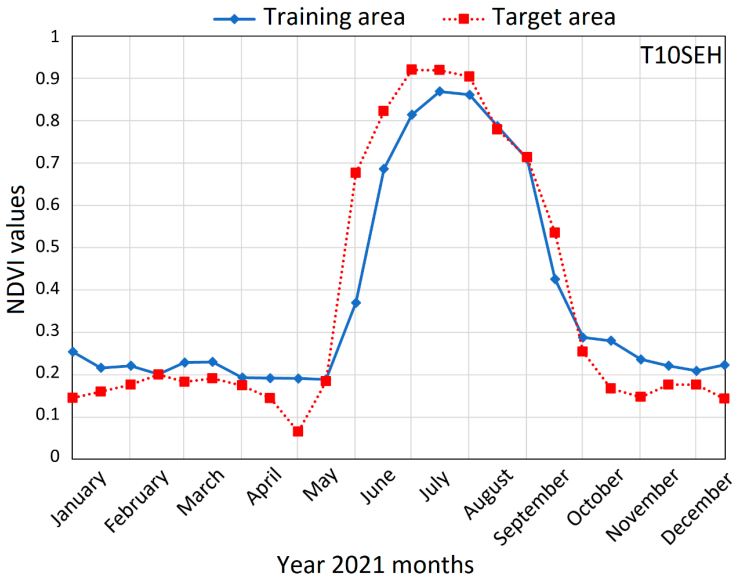


Figure 11. NDVI profiles of the training area (blue) and target area T10SEH (red) for rice.

The AMM and VIM methods identified June–August–September as the optimal three-month time series, demonstrating strong performance across various metrics, as shown in Table 3. The IMM method also produced effective results with the July–August–September combination, achieving an OA that was 0.71% lower compared to that achieved using the AMM method. Among the DL models trained on single-month data, August achieved an OA of 97.8%. While the IMM method demonstrated strong performance, visual interpretation identified August–September as the optimal time series, achieving an OA of 97.9%. Figure 12 illustrates a comparison of DL model performance for different time series combinations for rice segmentation in T10SEH.

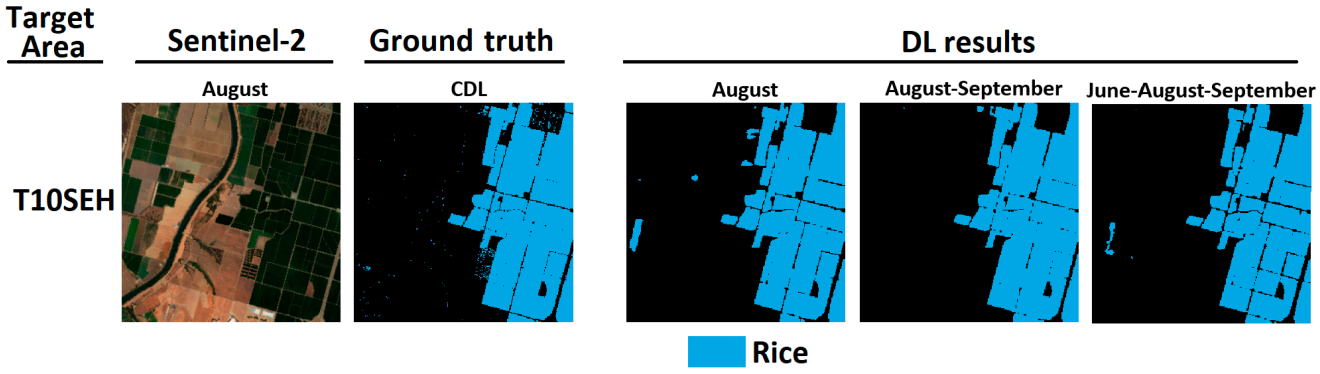


Figure 12. Visual comparison of results of transfer DL models for rice segmentation in T10SEH.

4. Discussion

4.1. Phenological Stage Alignment Methods

In this study, AMM proved to be a robust method for matching time series between the target and training areas in crop segmentation tasks, without the need for prior knowledge of cropland distributions across various target regions. By aligning NDVI profiles from Sentinel-2 images with the phenological stages of major crops, AMM effectively captured critical growth phases, improving segmentation accuracy by up to 20% in terms of OA compared to the STS method in some cases. While IMM demonstrated strong performance through a systematic identification of the best time series match using MCDA, its dependence on existing cropland knowledge limits its applicability in data-scarce scenarios. VIM, although practical and based on expert interpretation, lacks the systematic and adaptive strengths of AMM and IMM.

Figure 13 compares the F1-scores for the AMM, IMM, VIM, and STS methods. AMM demonstrated notable strength in real transfer learning scenarios, achieving comparable performance to IMM despite the absence of prior cropland knowledge. AMM improved F1-scores compared to the STS method by approximately 15% for corn, 6% for soybeans, 1% for rice, and 11% for other crops. It effectively captured critical growth phases of crops, leading to strong segmentation results for corn, soybeans, and rice. However, AMM faced challenges in areas like T17TNJ (Canada) due to distinct environmental conditions and differences in cropland definitions, highlighting the need for better dataset alignment. The STS method exhibited the weakest performance, with F1-scores in some cases more than 15% lower than AMM, highlighting the importance of methods like AMM for accurately capturing phenological stages in diverse target areas.

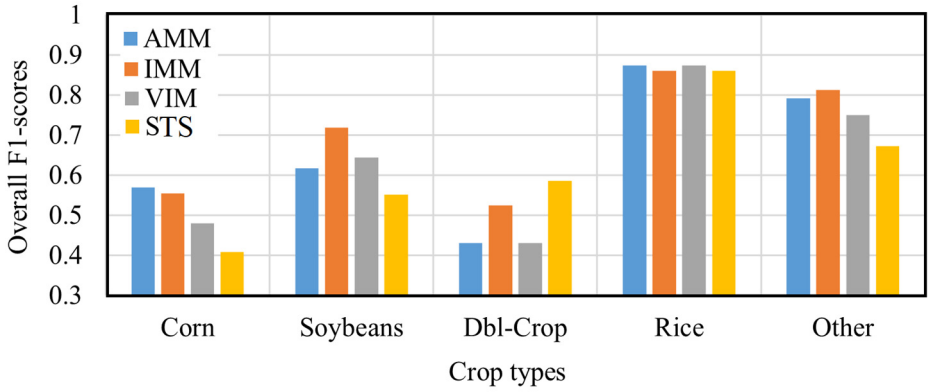


Figure 13. Comparison of average F1-scores across different crops (corn, soybeans, Dbl-Crop, rice, and other) for the AMM, IMM, VIM, and STS phenological stage alignment methods.

To compare AMM with SMF-based phenological alignment methods [28–30], the STSMF [30] method incorporates spatiotemporal information by using the VI time series of neighboring pixels from previous years, assuming that crop planting activities in local areas exhibit consistent patterns. However, in our scenario, this assumption is not practical due to limited knowledge about the target area. Specifically, while we may be aware of the major crops in the target region, we lack precise information about their locations and phenological characteristics. Furthermore, model-fitting approaches such as SMF-based methods [28–30] attempt to align an identical shape model with the VI time series of various target pixels. These methods rely on scaling and time shift operations to account for variations in the shape of the VI curves. However, scaling operations are highly sensitive to noise in the VI curves, which becomes a significant challenge in transfer learning scenarios. When dealing with incomplete or noisy VI time series, curve matching accuracy declines substantially. In contrast, the AMM method avoids exact NDVI curve matching

by identifying the optimal month-to-month correspondence between the training and target areas. This approach not only simplifies the process by requiring fewer parameters (limited to differences and slope calculations for three options per training month) but also eliminates the need for additional ground-measured phenological data or multi-year information from surrounding pixels, as required by STSMF.

Moreover, the AMM method is designed to accommodate large-scale applications with minimal computational effort. By averaging the NDVI values across large study areas, which may span 100 km by 100 km or more, AMM ensures scalability without compromising efficiency. This contrasts sharply with STSMF, which requires matching NDVI curves across multiple neighboring pixels, significantly increasing computational demands. Similarly, iterative methods like SMF-S are computationally expensive and impractical for large-area studies. Additionally, SMF-based methods often require extensive parameter tuning, further limiting their scalability. The AMM method, on the other hand, eliminates the need for parameter adjustments, making it not only computationally efficient but also more suitable for transfer learning in areas with limited prior information or irregular crop rotation patterns. These advantages highlight the superiority of the AMM method over SMF-based methods, particularly in scenarios where data availability and computational efficiency are critical.

4.2. One, Two-, and Three-Month Time Series

For the comparison of one-, two-, and three-month time series, we calculated the average values of F1-scores from different target areas. The charts in Figure 14 display these averages for the best results of each time series. The results did not show consistent trends across all combinations. In some cases, a single month yielded better performance for transfer learning, while in other cases, two or three months provided better results. For example, in the case of corn, increasing the number of months in the time series led to a decrease in F1-scores by approximately 12%. However, for soybeans, Dbl-Crop, and other classes, adding more months improved accuracy. This can be attributed to the more complex phenological stages of soybeans, Dbl-Crop, and other crops during their growth cycles, requiring a longer time series to capture these variations effectively.

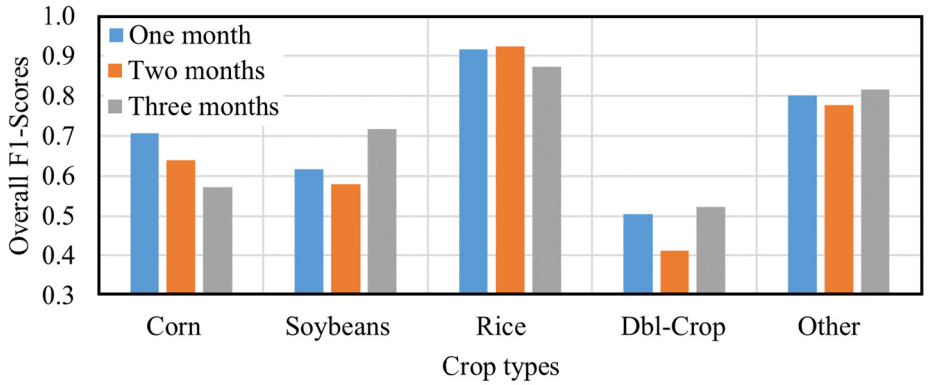


Figure 14. Comparison of average F1-scores for one-month, two-month, and three-month Sentinel-2 composite images across different crop types from the overall values of different target areas.

In contrast, for rice and corn, the phenological stages are less complex, and their growth can be adequately tracked with one- or two-month composite images. This finding highlights the importance of identifying specific time series for each crop, as well as ensuring proper time series matching between the training and target areas. In this study, we addressed this by grouping crops into three categories based on the major crops in the target areas. It may also be beneficial in future work to develop separate DL models for

each crop, trained on different datasets and time series, rather than using a single model for all crops.

4.3. Comparative Analysis with Prior Research

A study by Ge et al. [23] developed a transferable DL model for crop mapping based on phenological matching using CDL and Landsat data. The model was trained in the Mississippi River Delta in Arkansas, USA, tested across the Corn Belt of the US for corn classification, and transferred to California and Liaoning, China, for rice classification. This article is a suitable comparison for our study because it also addresses the challenge of transfer learning in crop classification across diverse geographic regions, focusing on the alignment of phenological stages between training and target areas. Additionally, the training and target data in their study are geographically close with slight differences, similar to our study, making the comparison relevant for evaluating the robustness and applicability of transfer learning methods in different environmental conditions.

Tables 4 and 5 provide detailed comparisons of the results we obtained using the AMM method against the results obtained by Ge et al. [23] using U-Net3. For instance, in this study, target area T14TNK is compared with T2 of the U-Net3 model. While T17TNJ (Canada) and T8 (China) are not geographically close, they serve as comparative examples outside of the CDL area. The U-Net3 was trained and evaluated using three images acquired from different regions on the same date over three consecutive periods [23]. The metrics used for comparison include OA, UA, PA, and F1-scores for corn across the eight target areas. In most target areas, including T14TNK, T14TNR, T15TUG, T16SCJ, and T16TFM, AMM effectively aligned the time series with the phenological stages of crops, achieving higher UA, PA, and F1-scores for corn, with F1-scores exceeding 50% in some cases.

Table 4. Comparison of segmentation results for corn and other crops achieved using AMM in our study and U-Net3 in Ge et al. [23]'s study for corresponding target areas from both studies (T1 to T8 for U-Net3 model).

Target Area	OA (%)	Corn [1] UA (%)	Corn [1] PA (%)	Corn [1] F1	Other UA (%)	Other PA (%)	Other F1
T14TNK (AMM)	72.91	94.80	49.60	0.63	73.11	82.68	0.77
T2 (U-Net3)	80.56	87.13	0.83	0.02	83.54	99.80	0.91
T14TNR (AMM)	64.82	95.70	52.10	0.67	66.76	58.62	0.62
T1 (U-Net3)	78.00	81.17	1.99	0.04	87.43	99.68	0.93
T15TUG (AMM)	80.11	83.41	92.27	0.87	73.17	66.24	0.69
T3 (U-Net3)	59.88	76.68	4.58	0.09	79.84	98.93	0.88
T16SCF (AMM)	90.00	93.87	76.00	0.84	91.72	95.26	0.93
T6 (U-Net3)	94.60	95.31	89.18	0.92	95.18	96.70	0.96
T16SCJ (AMM)	75.94	98.10	48.99	0.64	63.36	95.32	0.75
T5 (U-Net3)	79.07	87.36	40.82	0.56	89.55	97.02	0.93
T16TFM (AMM)	76.91	95.73	27.51	0.42	74.64	97.91	0.85
T4 (U-Net3)	77.72	86.76	12.71	0.22	84.13	99.31	0.91
T17TNJ (Canada)	76.92	95.62	8.54	0.16	77.12	99.28	0.87
T8 (China, U-Net3)	83.15	43.21	88.61	0.58	97.13	77.73	0.86

[1] Results of corn and other are separated with distinct backgrounds.

Table 5. Comparison of segmentation results for rice and other crops in T10SEH target area with T7 in Ge et al. [23]'s study.

Target Area	OA (%)	Rice [1] UA (%)	Rice [1] PA (%)	Rice [1] F1	Other UA (%)	Other PA (%)	Other F1
T10SEH (AMM)	97.88	85.36	92.83	0.87	98.96	98.47	0.98
T7 (U-Net3)	94.31	90.88	91.33	0.91	95.25	95.03	0.95

[1] Results of rice and other are separated with distinct backgrounds.

Additionally, Table 5 shows that for rice in the T10SEH target area, AMM achieved a better OA of 97.88%, compared to 94.31% OA achieved using the U-Net3 method, further underscoring its efficacy. The AMM method's ability to adapt to specific growth phases of crops, producing reliable segmentation outcomes across diverse agricultural settings, highlights its effectiveness and potential for broader application.

Despite the notable strengths of the AMM method, several weaknesses are observed when comparing accuracy metrics with U-Net3. The OA and F1-scores for corn in target areas such as T16SCF and T17TNJ (Canada) in our study are significantly lower than those achieved by the U-Net3 method, with a difference of up to 8% in corn's F1-score for T16SCF. One major factor is that Ge's study did not include soybeans, focusing solely on corn, which likely led to more optimized models, whereas our models had to balance classification among multiple crops, potentially diluting accuracy. In contrast, our results for rice, where the model was trained on a single crop, were comparable with their findings. Additionally, our extensive preprocessing efforts to exclude no-data regions, such as areas affected by clouds and satellite image edges, might have reduced apparent accuracy for the other class, whereas Ge et al.'s study might have inflated accuracy metrics without similar preprocessing. Furthermore, differences in the number of training samples and the specific characteristics of the geographical areas studied, along with the size and focus of the training area and the systematic tuning of hyperparameters, all impact model performance.

4.4. Implementation Prospects and Limitations

The AMM method introduced in this study shows great potential for improving transfer crop segmentation DL models by aligning phenological stages between training and target areas. Its adaptability allows for effective performance across regions without requiring detailed prior knowledge of local crop distributions. By aligning NDVI profiles, AMM accurately captures crop growth stages, improving segmentation accuracy across diverse locations, with up to a 20% increase in OA compared to the STS method in some cases. This makes AMM a cost-effective solution for large-scale agricultural applications, such as crop health monitoring and yield estimation, while reducing the need for extensive local training data.

Despite its advantages, the AMM method has several limitations. One significant limitation is the use of monthly composites for data reduction, which restricts the ability to closely track phenological processes. Employing bi-weekly data composites or integrating multiple satellite datasets, such as Sentinel-2 and Landsat, could enhance temporal resolution and provide a more detailed representation of crop growth stages. Additionally, the method assumes consistent phenological stages for crops across different regions; however, geographic, climatic, and environmental variations can significantly impact crop phenology. Future research should consider incorporating location-specific data, such as local weather patterns, crop varieties, and soil characteristics, to better account for regional differences in crop growth stages. In AMM, only two constraints were used for time series matching, but incorporating additional constraints, such as soil moisture levels, temperature variations, or additional vegetation indices, could further improve alignment accuracy.

5. Conclusions

In this study, we introduced and evaluated the AMM method to improve transfer learning in crop segmentation tasks by aligning NDVI profiles between training and target areas. AMM addressed phenological differences in crop growth stages that typically hinder the transferability of DL models. Using the GEE to process large Sentinel-2 datasets, we optimized time series data for major crops like corn, soybeans, rice, and Dbl-Crop. Our experiments across diverse regions in the US and Canada demonstrated AMM's

effectiveness, improving segmentation accuracy by up to 20% in terms of OA compared to using the same time series as the time series of the training area. The method consistently performed well, particularly in rice-growing regions, achieving an OA of 97.88% and often matching or exceeding other phenological matching techniques. It also often matched or exceeded other phenological matching techniques in corn segmentation, with an average OA across all target areas exceeding 79% for cropland segmentation. This study contributes to enhancing the accuracy of crop segmentation in diverse environments, supporting more effective agricultural monitoring and management. Future work will focus on further refining the AMM method, exploring additional crop types, and extending its application to other geographic regions while accounting for location-based variability in crop phenology to enhance its generalizability and impact.

Author Contributions: R.M.: conceptualization, methodology, formal analysis, validation, investigation, data curation, visualization, writing—original draft; F.W.: conceptualization, methodology, validation, resources, writing—review and editing, supervision, projection administration, funding acquisition; G.Q.: supervision, validation, writing—review and editing, funding acquisition; A.O.: validation, investigation, writing—review and editing; L.F.: software, validation, writing—review and editing; G.Y.: writing—review and editing, supervision, project administration. All authors have read and agreed to the published version of the manuscript.

Funding: This study was conducted without any external financial support.

Data Availability Statement: Sentinel-2 imagery is freely accessible through the ESA Copernicus Data Space Ecosystem (https://dataspace.copernicus.eu/), and CDL data are available from the USDA NASS website (https://www.nass.usda.gov). ACI data can be accessed through Agriculture and Agri-Food Canada's (AAFC) website (https://agriculture.canada.ca/). The dataset is available on reasonable request due to the terms and conditions of data sharing agreements. Also, Sentinel-2 composite was processed in the GEE environment (https://earthengine.google.com).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, X.; Chen, N.; Sheng, H.; Ip, C.; Yang, L.; Chen, Y.; Sang, Z.; Tadesse, T.; Lim, T.P.Y.; Rajabifard, A.; et al. Urban drought challenge to 2030 sustainable development goals. *Sci. Total Environ.* **2019**, *693*, 133536. [CrossRef] [PubMed]
2. Wang, Z.; Sun, X.; Liu, X.; Xu, F.; Huang, H.; Ti, R.; Yu, H.; Wang, Y.; Wei, Y. Improved Paddy Rice Classification Utilizing Sentinel-1/2 Imagery in Anhui China: Phenological Features, Algorithms, Validation and Analysis. *Agriculture* **2024**, *14*, 1282. [CrossRef]
3. Zhao, B.; Huang, B.; Zhong, Y. Transfer Learning with Fully Pretrained Deep Convolution Networks for Land-Use Classification. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1436–1440. [CrossRef]
4. Yang, L.; Huang, R.; Zhang, J.; Huang, J.; Wang, L.; Dong, J.; Shao, J. Inter-Continental Transfer of Pre-Trained Deep Learning Rice Mapping Model and Its Generalization Ability. *Remote Sens.* **2023**, *15*, 2443. [CrossRef]
5. Song, W.; Feng, A.; Wang, G.; Zhang, Q.; Dai, W.; Wei, X.; Hu, Y.; Amankwah, S.O.Y.; Zhou, F.; Liu, Y. Bi-Objective Crop Mapping from Sentinel-2 Images Based on Multiple Deep Learning Networks. *Remote Sens.* **2023**, *15*, 3417. [CrossRef]
6. Moghimi, A.; Khodadadzadeh, M.; Li, J.; Plaza, A.J. Review of deep learning for remote sensing data. *Remote Sens.* **2018**, *10*, 638. [CrossRef]
7. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep Learning in Agriculture: A Survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]
8. Viana, C.M.; Freire, D.; Abrantes, P.; Rocha, J.; Pereira, P. Agricultural land systems importance for supporting food security and sustainable development goals: A systematic review. *Sci. Total Environ.* **2022**, *806*, 150718. [CrossRef]
9. Csikós, N.; Szabó, B.; Hermann, T.; Laborczi, A.; Matus, J.; Pásztor, L.; Szatmári, G.; Takács, K.; Tóth, G. Cropland Productivity Evaluation: A 100 m Resolution Country Assessment Combining Earth Observation and Direct Measurements. *Remote Sens.* **2023**, *15*, 1236. [CrossRef]
10. Ahirwal, J.; Sahoo, U.K.; Thangjam, U.; Thong, P. Oil palm agroforestry enhances crop yield and ecosystem carbon stock in northeast India: Implications for the United Nations sustainable development goals. *Sustain. Prod. Consum.* **2022**, *30*, 478–487. [CrossRef]

11. United Nations Development Program, U.N.D.P. Sustainable Development Goals. Available online: https://www.undp.org/sustainable-development-goals (accessed on 20 July 2024).
12. Htitiou, A.; Boudhar, A.; Lebrini, Y.; Hadria, R.; Lionboui, H.; Benabdelouahab, T. A comparative analysis of different phenological information retrieved from Sentinel-2 time series images to improve crop classification: A machine learning approach. *Geocarto Int.* **2022**, *37*, 1426–1449. [CrossRef]
13. Lark, T.J.; Schelly, I.H.; Gibbs, H.K. Accuracy, Bias, and Improvements in Mapping Crops and Cropland across the United States Using the USDA Cropland Data Layer. *Remote Sens.* **2021**, *13*, 968. [CrossRef]
14. Hao, P.; Di, L.; Zhang, C.; Guo, L. Transfer Learning for Crop classification with Cropland Data Layer data (CDL) as training samples. *Sci. Total Environ.* **2020**, *733*, 138869. [CrossRef] [PubMed]
15. Sun, L.; Lou, Y.; Shi, Q.; Zhang, L. Spatial domain transfer: Cross-regional paddy rice mapping with a few samples based on Sentinel-1 and Sentinel-2 data on GEE. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *128*, 103762. [CrossRef]
16. You, N.; Dong, J. Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 109–123. [CrossRef]
17. Hoppe, H.; Dietrich, P.; Marzahn, P.; Weiß, T.; Nitzsche, C.; Freiherr von Lukas, U.; Wengerek, T.; Borg, E. Transferability of Machine Learning Models for Crop Classification in Remote Sensing Imagery Using a New Test Methodology: A Study on Phenological, Temporal, and Spatial Influences. *Remote Sens.* **2024**, *16*, 1493. [CrossRef]
18. Shi, Y.; Li, L.; Wu, B.; Zhang, Y.; Wang, B.; Niu, W.; He, L.; Jin, N.; Pan, S.; Tian, H.; et al. Predicting rice productivity for ground data-sparse regions: A transferable framework and its application to North Korea. *Sci. Total Environ.* **2024**, *946*, 174227. [CrossRef] [PubMed]
19. Antonijević, O.; Jelić, S.; Bajat, B.; Kilibarda, M. Transfer learning approach based on satellite image time series for the crop classification problem. *J. Big Data* **2023**, *10*, 54. [CrossRef]
20. Xu, J.; Zhu, Y.; Zhong, R.; Lin, Z.; Xu, J.; Jiang, H.; Huang, J.; Li, H.; Lin, T. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. *Remote Sens. Environ.* **2020**, *247*, 111946. [CrossRef]
21. Chen, S.; Yi, Q.; Wang, F.; Zheng, J.; Li, J. Improving the matching degree between remotely sensed phenological dates and physiological growing stages of soybean by a dynamic offset-adjustment strategy. *Sci. Total Environ.* **2024**, *906*, 167783. [CrossRef]
22. Shen, Y.; Zhang, X.; Yang, Z. Mapping corn and soybean phenometrics at field scales over the United States Corn Belt by fusing time series of Landsat 8 and Sentinel-2 data with VIIRS data. *ISPRS J. Photogramm. Remote Sens.* **2022**, *186*, 55–69. [CrossRef]
23. Ge, S.; Zhang, J.; Pan, Y.; Yang, Z.; Zhu, S. Transferable deep learning model based on the phenological matching principle for mapping crop extent. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102451. [CrossRef]
24. Yang, G.; Li, X.; Liu, P.; Yao, X.; Zhu, Y.; Cao, W.; Cheng, T. Automated in-season mapping of winter wheat in China with training data generation and model transfer. *ISPRS J. Photogramm. Remote Sens.* **2023**, *202*, 422–438. [CrossRef]
25. Diao, C. Remote sensing phenological monitoring framework to characterize corn and soybean physiological growing stages. *Remote Sens. Environ.* **2020**, *248*, 111960. [CrossRef]
26. Segarra, J.; Araus, J.L.; Kefauver, S.C. Farming and Earth Observation: Sentinel-2 data to estimate within-field wheat grain yield. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *107*, 102697. [CrossRef]
27. Yang, Z.; Diao, C.; Gao, F.; Li, B. EMET: An emergence-based thermal phenological framework for near real-time crop type mapping. *ISPRS J. Photogramm. Remote Sens.* **2024**, *215*, 271–291. [CrossRef]
28. Sakamoto, T.; Wardlow, B.D.; Gitelson, A.A.; Verma, S.B.; Suyker, A.E.; Arkebauer, T.J. A Two-Step Filtering approach for detecting maize and soybean phenology with time-series MODIS data. *Remote Sens. Environ.* **2010**, *114*, 2146–2159. [CrossRef]
29. Liu, L.; Cao, R.; Chen, J.; Shen, M.; Wang, S.; Zhou, J.; He, B. Detecting crop phenology from vegetation index time-series data by improved shape model fitting in each phenological stage. *Remote Sens. Environ.* **2022**, *277*, 113060. [CrossRef]
30. Cao, R.; Li, L.; Liu, L.; Liang, H.; Zhu, X.; Shen, M.; Zhou, J.; Li, Y.; Chen, J. A spatiotemporal shape model fitting method for within-season crop phenology detection. *ISPRS J. Photogramm. Remote Sens.* **2024**, *217*, 179–198. [CrossRef]
31. Hamrouni, Y.; Paillassa, E.; Chéret, V.; Monteil, C.; Sheeren, D. From local to global: A transfer learning-based approach for mapping poplar plantations at national scale using Sentinel-2. *ISPRS J. Photogramm. Remote Sens.* **2021**, *171*, 76–100. [CrossRef]
32. Konduri, V.S.; Kumar, J.; Hargrove, W.W.; Hoffman, F.M.; Ganguly, A.R. Mapping crops within the growing season across the United States. *Remote Sens. Environ.* **2020**, *251*, 112048. [CrossRef]
33. Zhang, C.; Di, L.; Lin, L.; Li, H.; Guo, L.; Yang, Z.; Yu, E.G.; Di, Y.; Yang, A. Towards automation of in-season crop type mapping using spatiotemporal crop information and remote sensing data. *Agric. Syst.* **2022**, *201*, 103462. [CrossRef]
34. Maleki, R.; Wu, F.; Oubara, A.; Fathollahi, L.; Yang, G. Refinement of Cropland Data Layer with Effective Confidence Layer Interval and Image Filtering. *Agriculture* **2024**, *14*, 1285. [CrossRef]
35. Abernethy, J.; Beeson, P.; Boryan, C.; Hunt, K.; Sartore, L. Preseason crop type prediction using crop sequence boundaries. *Comput. Electron. Agric.* **2023**, *208*, 107768. [CrossRef]

36. Tran, K.H.; Zhang, H.K.; McMaine, J.T.; Zhang, X.; Luo, D. 10 m crop type mapping using Sentinel-2 reflectance and 30 m cropland data layer product. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *107*, 102692. [CrossRef]
37. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
38. Wang, S.; Azzari, G.; Lobell, D.B. Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. *Remote Sens. Environ.* **2019**, *222*, 303–317. [CrossRef]
39. Massey, R.; Sankey, T.T.; Yadav, K.; Congalton, R.G.; Tilton, J.C. Integrating cloud-based workflows in continental-scale cropland extent classification. *Remote Sens. Environ.* **2018**, *219*, 162–179. [CrossRef]
40. Pouliot, D.; Alavi, N.; Wilson, S.; Duffe, J.; Pasher, J.; Davidson, A.; Daneshfar, B.; Lindsay, E. Assessment of Landsat Based Deep-Learning Membership Analysis for Development of from–to Change Time Series in the Prairie Region of Canada from 1984 to 2018. *Remote Sens.* **2021**, *13*, 634. [CrossRef]
41. Onojeghuo, A.O.; Miao, Y.; Blackburn, G.A. Deep ResU-Net Convolutional Neural Networks Segmentation for Smallholder Paddy Rice Mapping Using Sentinel 1 SAR and Sentinel 2 Optical Imagery. *Remote Sens.* **2023**, *15*, 1517. [CrossRef]
42. Ferla, G.; Mura, B.; Falasco, S.; Caputo, P.; Matarazzo, A. Multi-Criteria Decision Analysis (MCDA) for sustainability assessment in food sector. A systematic literature review on methods, indicators and tools. *Sci. Total Environ.* **2024**, *946*, 174235. [CrossRef] [PubMed]
43. ESA. Copernicus Sentinel-2 Collection 1 MSI Level-2A (L2A). Available online: https://sentinels.copernicus.eu/web/sentinel/sentinel-data-access/sentinel-products/sentinel-2-data-products/collection-1-level-2a (accessed on 17 July 2024).
44. Oubara, A.; Wu, F.; Maleki, R.; Ma, B.; Amamra, A.; Yang, G. Enhancing Adversarial Learning-Based Change Detection in Imbalanced Datasets Using Artificial Image Generation and Attention Mechanism. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 125. [CrossRef]
45. Li, Z.; Deng, X.; Lan, Y.; Liu, C.; Qing, J. Fruit tree canopy segmentation from UAV orthophoto maps based on a lightweight improved U-Net. *Comput. Electron. Agric.* **2024**, *217*, 108538. [CrossRef]
46. Mckay, M.; Danilevicz, M.F.; Ashworth, M.B.; Rocha, R.L.; Upadhyaya, S.R.; Bennamoun, M.; Edwards, D. Focus on the Crop Not the Weed: Canola Identification for Precision Weed Management Using Deep Learning. *Remote Sens.* **2024**, *16*, 2041. [CrossRef]
47. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; CRC Press: Boca Raton, FL, USA, 2019.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.